

Studienarbeit
Kalibrierung einer Fischaugenkamera

Adrian Posor

WS 2004/05

Betreuer: Jan Woetzel

Arbeitsgruppe Multimediale Informationsverarbeitung,
Prof. Dr. Ing. R. Koch
Institut für Informatik und Praktische Mathematik
Christian-Albrechts-Universität zu Kiel

Zusammenfassung

Um mit einer Kamera möglichst genaue Messungen vornehmen zu können, muß zu jedem Bildpunkt der zugehörige Strahl im dreidimensionalen Raum bekannt sein. Aufgrund von Toleranzen bei der Fertigung muß diese Abbildungsfunktion für jede zu verwendende Optik extra bestimmt werden. Dies gilt auch für Kameras, die keine Fischaugenkameras sind, doch sind im Fall von Fischaugenkameras einige wesentliche Unterschiede zu beachten, die in dieser Arbeit erläutert werden.

In dieser Studienarbeit wird ein Verfahren zur Berechnung der Funktion vorgestellt, die die Abweichung der Abbildung der Kamera von der einer idealen Kamera beschreibt.

Inhaltsverzeichnis

1	Einleitung.....	3
2	Grundlagen.....	4
2.1	Kameramodelle.....	4
2.1.1	Lochkameramodell.....	4
2.1.2	Modell einer idealen Kugelkamera.....	5
2.2	Darstellung von Fischaugenkamerabildern.....	6
2.3	Kamerakalibrierung.....	8
2.3.1	Ideales Modell.....	8
2.3.2	Abweichung vom idealen Modell.....	9
2.3.3	Approximation des Modells.....	9
2.3.4	Ursachen für Verzerrungen.....	9
2.4	Canny Kantendetektor.....	10
3	Berechnung einer Entzerrungsfunktion.....	11
3.1	Überblick über das Verfahren.....	11
3.2	Bestimmung der Linien des Kalibrierungsmusters.....	13
3.3	Bestimmung des optischen Zentrums.....	14
3.4	Bestimmung der radialsymmetrischen Entzerrungsfunktion.....	15
4	Ergebnisse und Bewertung.....	17
4.1	Ergebnisse.....	17
4.2	Bewertung.....	28
4.3	Ausblick.....	29
5	Literatur.....	29

1 Einleitung

Mit Hilfe von Kameras lassen sich Messungen durchführen (z.B. Messen von Entfernungen, Erstellen von Tiefenkarten). Damit diese Messungen möglichst genau sind, muß die mathematische Funktion bekannt sein, mit der sich berechnen läßt, welcher Lichtstrahl im dreidimensionalen Raum zu welchem Punkt in der Bildebene der Kamera abgebildet wird. Aufgrund von Toleranzen bei der Produktion der Kameras sind die Abbildungseigenschaften jeder Kamera etwas anders, weshalb diese Funktion für jede Kamera einzeln bestimmt werden muß. Die Beschreibung der Abbildung von Lichtstrahlen zu Bildpunkten auf der Bildebene der Kamera erfolgt durch ein Kameramodell. Dieses ist beschrieben durch interne und externe Kameraparameter. Zu den inneren Kameraparametern gehören z.B. die Brennweite sowie die Lage des optischen Zentrums der Kamera. Zu den externen Kameraparametern gehören die Position und Lage der Kamera im Weltkoordinatensystem. Die Bestimmung dieser Kameraparameter nennt man Kamerakalibrierung.

Konventionelle Kameras haben einen sehr beschränkten Blickwinkel. Im Gegensatz dazu haben sogenannte Fischaugenkameras einen Blickwinkel von über 180 Grad. Bei konventionellen Kameras kann das Lochkameramodell zur Beschreibung des Strahlenganges in der Kamera verwendet werden. Für Fischaugenkameras muß ein angepaßtes Modell verwendet werden.

Es gibt bereits verschiedene Kamerakalibrierungsverfahren, sowohl für konventionelle als auch für Fischaugenkameras, doch arbeiten die meisten von ihnen mit dreidimensionalen Kalibrierungsobjekten oder mit Bildsequenzen. Bei Verwendung eines dreidimensionalen Kalibrierungsobjektes kennt man die 3D-Koordinaten dieses Objektes und kann daher auf der Grundlage eines Kameramodells und des von der Kamera aufgenommenen Bildes alle Kameraparameter (wie z.B. Brennweite, Position des optischen Zentrums) bestimmen. Das hier vorgestellte Verfahren ist für Fischaugenkameras geeignet und benötigt nur ein eindimensionales Kalibrierungsmuster. Es kann mit einem einzigen Bild auskommen, doch werden mit Bildsequenzen bessere Ergebnisse erzielt. Das Verfahren macht sich zu Nutze, daß gerade Linien durch die Projektion nicht in ihrer Form verändert werden, sofern die Kamera nicht ein verzerrtes Bild liefert.

In Kapitel 2.1 werden die hier verwendeten Kameramodelle beschrieben. Es folgt eine Vorstellung verschiedener Darstellungsformen von Fischaugenkamerabildern in Abschnitt 2.2. Anschließend werden die Grundlagen der Kamerakalibrierung erörtert. Es werden verschiedene Modelle vorgestellt und die Ursachen für Verzerrungen aufgezeigt. In Kapitel 3 wird erklärt, auf welche Weise in dieser Arbeit die Berechnung einer Entzerrungsfunktion für Kamerabilder durchgeführt wird. Angefangen bei der Erkennung der Linien des Kalibrierungsmusters über die Bestimmung des optischen Zentrums bis hin zur eigentlichen Berechnung der radialsymmetrischen Entzerrungsfunktion mithilfe der Methode der kleinsten Quadrate werden alle wichtigen Details aufgegriffen. Abschließend werden Ergebnisse in Form von entzerrten Bildern und errechneten Entzerrungsfunktionen präsentiert.

2 Grundlagen

2.1 Kameramodelle

2.1.1 Lochkameramodell

Zur Beschreibung konventioneller Kameras eignet sich das Lochkameramodell. Die Abbildung von Punkten $M=(X, Y, Z)$ im dreidimensionalen Raum zu den 2D-Punkten im Bild der Kamera $m=(x, y)$ ist in diesem Modell wie folgt:

$$x = \frac{f}{Z} \cdot X \quad \text{und} \quad y = \frac{f}{Z} \cdot Y$$

mit f = Brennweite der Kamera.

In Kugelkoordinaten ausgedrückt bedeutet dies:

Es seien $M=(R, \Phi, \Theta)$ und $m=(r, \phi)$. Dann gilt:

$$r = f \cdot \tan \Theta \quad \text{und} \quad \phi = \Phi$$

In diesem Kameramodell ist ein Blickwinkel von 180 Grad oder mehr problematisch. Punkte mit $\Theta = \pm 90^\circ$ liegen im Unendlichen und für Punkte mit $|\Theta| > 90^\circ$ kann man nicht entscheiden, ob sie vor oder hinter der Bildebene liegen. Dies zeigt sich auch in den obigen Gleichungen. Für $Z=0$ bzw. $\Theta=90^\circ$ sind sie nicht definiert. Punkte mit negativer z-Koordinate schneiden die Bildebene gar nicht bzw. man erhält den Schnittpunkt mit der Bildebene der sich durch Punktspiegelung am Zentrum C ergibt. Da Fischaugenkameras diese Einschränkung nicht besitzen, braucht man für ihre Beschreibung ein allgemeineres Modell.

Die folgende Abbildung zeigt das Lochkameramodell mit den hier verwendeten Bezeichnungen. C bezeichnet das optische Zentrum der Kamera. Es fällt in der hier gezeigten Abbildung mit dem Ursprung des Weltkoordinatensystems zusammen. Daher entspricht die Z-Achse hier der optischen Achse der Kamera. Dies ist im allgemeinen nicht der Fall, weshalb wir das optische Zentrum erst noch bestimmen müssen.

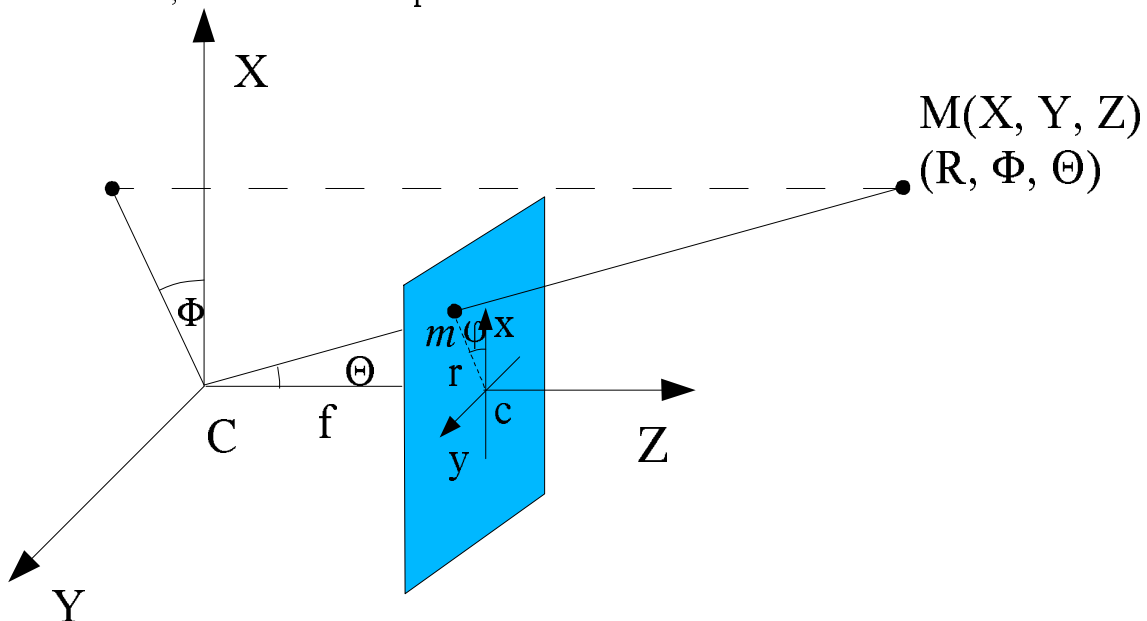


Abbildung 1: Lochkameramodell. Das optische Zentrum der Kamera fällt in dieser Darstellung mit dem Weltkoordinatensystem zusammen.

2.1.2 Modell einer idealen Kugelkamera

Für Fischaugenkameras gelten andere Projektionsgleichungen. Im folgenden sei der Winkel zwischen z-Achse und Lichtstrahl mit θ bezeichnet anstatt mit Θ . Für alle diese Kameras gilt zwar $\phi = \Phi$, aber $r = f \cdot \theta$ für *equidistante Projektion*, bei der in dieser Studienarbeit verwendeten Fischaugenkamera. Das hier dargestellte Verfahren soll die Abbildungsfunktion noch genauer bestimmen, als sie durch diese Gleichung bereits beschrieben wird.

Abbildung 2 zeigt das Fischaugenkameramodell. Der Punkt P wird auf den Punkt p abgebildet. Zum Vergleich: Eine Lochkamera bildet P auf p' ab.

Abbildung 3 zeigt den Strahlengang einer Fischaugenkamera.

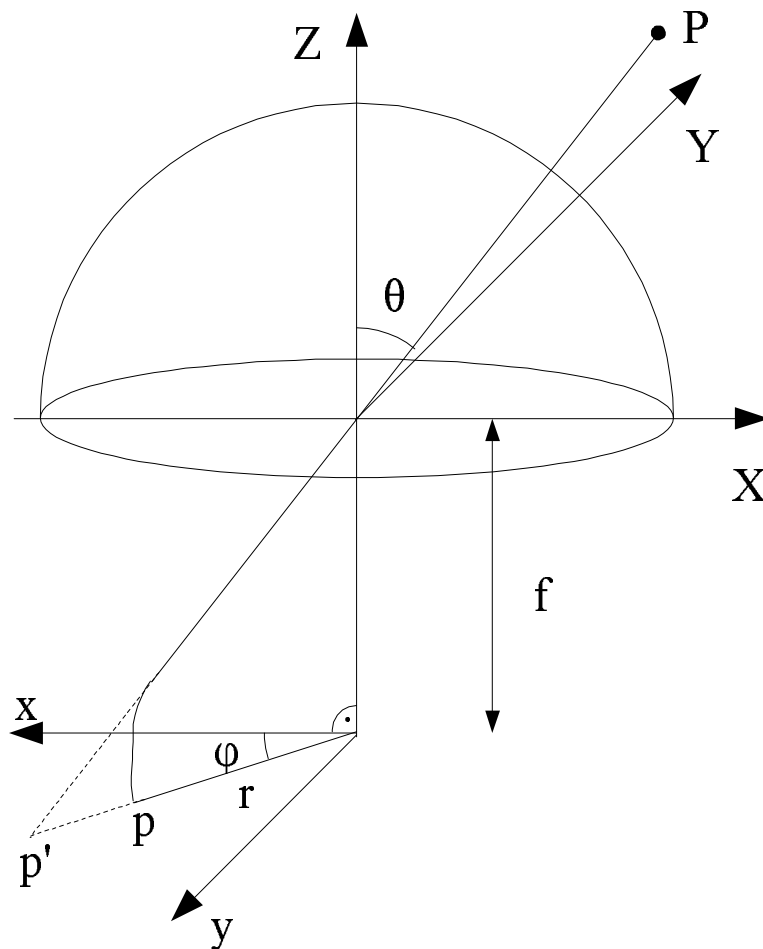


Abbildung 2: Modell der idealen Kugelkamera

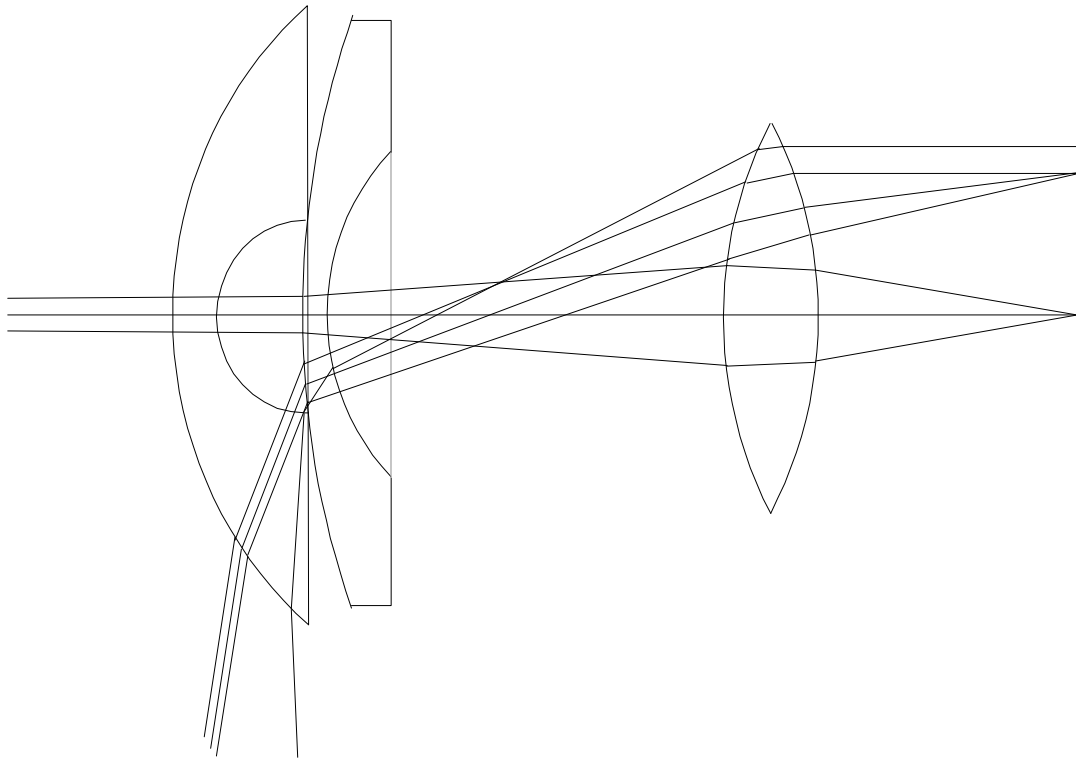


Abbildung 3: Strahlengang einer Fischaugenkamera

2.2 Darstellung von Fischaugenkamerabildern

Von Fischaugenkameras aufgenommene Bilder lassen sich auch auf andere Weise darstellen. Jeder Punkt im Bild lässt sich sowohl durch Angabe der x- und y-Koordinaten darstellen als auch in Kugelkoordinaten, also θ , ϕ und r . Bei der Projektion gehen die Tiefeninformationen verloren. In Kugelkoordinaten fällt durch die Projektion der Radius r weg. Die folgende Abbildung zeigt die Beschreibung von Punkten im Raum durch Kugelkoordinaten.

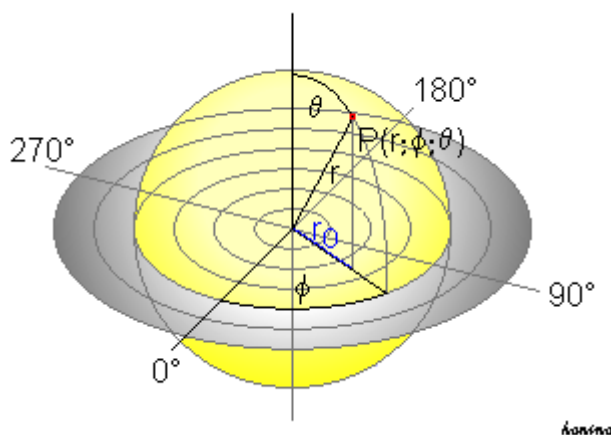


Abbildung 4: Kugelkoordinaten, entnommen aus [5]

Trägt man θ auf der x-Achse und ϕ auf der y-Achse auf, so erhält man eine Abwicklung des Bildes. Kreise um das Zentrum im Kamerabild werden zu Linien im Theta-Phi-Bild. Abbildung 5 zeigt skizzenhaft den Zusammenhang zwischen der Darstellung eines Bildes mit den Koordinaten x und y und der Darstellung mit den Koordinaten θ und ϕ .

Abbildung 6 zeigt ein Beispiel einer solchen Abwicklung anhand eines mit der hier verwendeten Fischaugenkamera aufgenommenen Bildes.

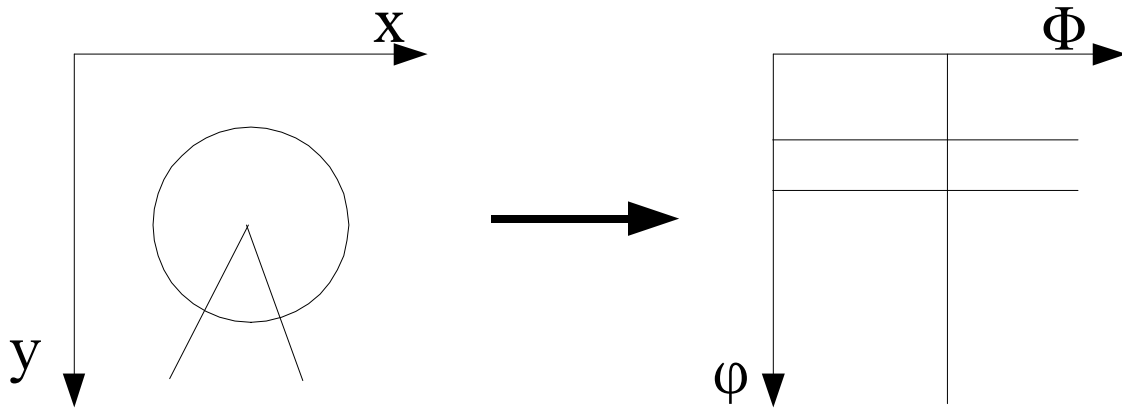


Abbildung 5: Umwandlung eines Bildes vom x/y -Koordinatensystem zum θ/φ -Koordinatensystem

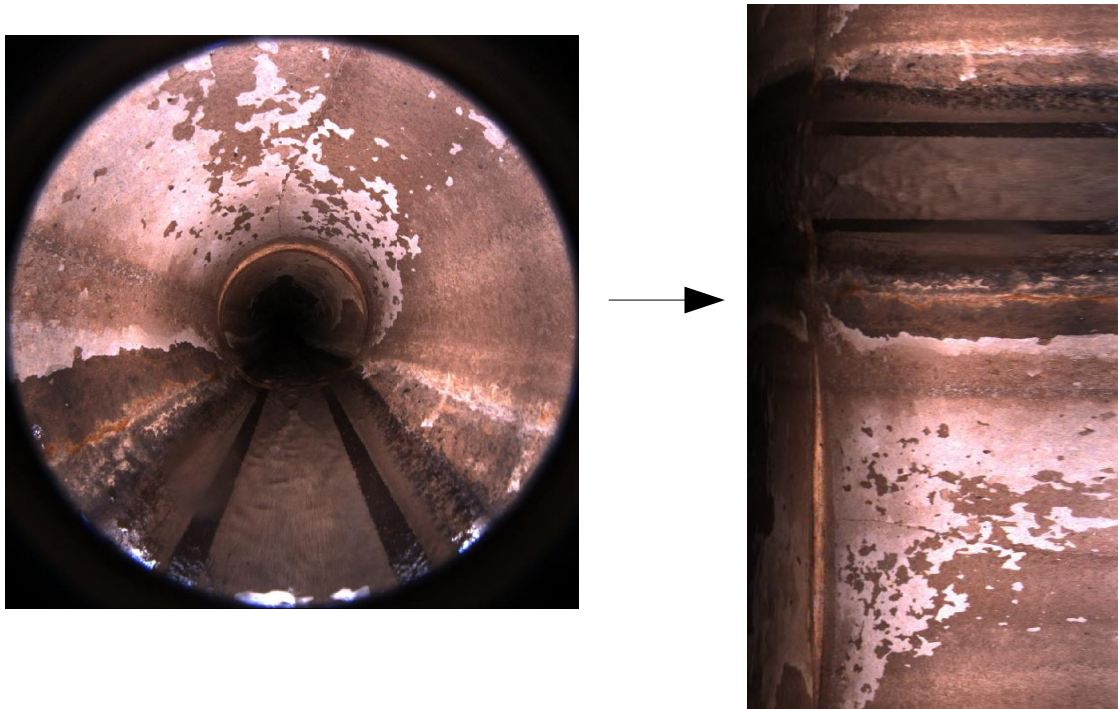


Abbildung 6: Umwandlung eines von einer Fischaugenkamera aufgenommenen Bildes vom x/y -Koordinatensystem zum θ/φ -Koordinatensystem

Eine weitere mögliche Darstellung des Fischaugenkamerabildes ist die Darstellung aus der Sicht einer Lochkamera. Dabei ist zu beachten, daß nicht das gesamte Bild in diese Form gebracht werden kann, da eine Lochkamera immer einen Blickwinkel von weniger als 180° hat. Man kann daher nur einen Ausschnitt von weniger als 180° aus dem Bild als Lochkamerabild zeigen. Die Blickrichtung des Ausschnittes kann beliebig gewählt werden. Mit mehreren Ausschnitten läßt sich das gesamte Bild darstellen. Abbildung 7 zeigt das in Abbildung 6 gezeigte Bild aus der Sicht einer Lochkamera mit einem Blickwinkel von 150° .



Abbildung 7: Von einer Fischaugenkamera aufgenommenes Bild, welches in die Sicht einer Lochkamera umgerechnet wurde.

2.3 Kamerakalibrierung

Ziel der Kamerakalibrierung ist es, die Parameter der Kamera entsprechend dem verwendeten Kameramodell finden. Durch die Parameter wird die Abbildungsfunktion der Kamera beschrieben.

Man beachte, daß gerade Linien unabhängig von ihrer Lage im dreidimensionalen Raum immer als gerade Linien im Kamerabild erscheinen. Krumme Linien im Kamerabild, von denen wir wissen, daß sie im dreidimensionalen Raum gerade sind, erhalten also ihre Krümmung ausschließlich von der Verzerrung, die die Kamera verursacht. Diese Verzerrung stellt eine Abweichung vom Lochkameramodell dar. Das verwendete Kalibrierungsmuster ist ein Schachbrettmuster von dem wir wissen, daß alle Linien, die durch die Felder des Schachbrettmusters gebildet werden, gerade sind.

2.3.1 Ideales Modell

In unserem Fall gehen wir davon aus, daß die hier verwendete Kamera der Gleichung

$$r = f \cdot \theta$$

genügt. Das bedeutet, daß alle Lichtstrahlen mit dem Winkel θ zur z-Achse den Abstand r vom optischen Zentrum im Bild haben. Die Kamera wird jedoch aufgrund von Ungenauigkeiten bei der Produktion von diesem Ideal abweichen. Um genau berechnen zu können, welcher Bildpunkt zu welchem Lichtstrahl gehört, müssen wir die Abweichung vom Ideal oder eine Approximation der Abbildungsfunktion berechnen.

2.3.2 Abweichung vom idealen Modell

Bilder, die mit einer idealen Kugelkamera aufgenommen wurden bezeichnen wir als unverzerrt, solche, die mit einer nicht idealen Kugelkamera aufgenommen wurden als verzerrt. Entsprechend bezeichnen wir Abstände vom Zentrum in einem unverzerrten Bild als r_u und Abstände vom Zentrum in einem verzerrten Bild als r_v .

Es gilt:

$$r_u = f \cdot \theta$$

und

$$r_v = r_u + \Delta = f \cdot \theta + \Delta$$

Daraus folgt:

$$r_u = r_v - \Delta = \kappa_1 \cdot r_v + \kappa_2 \cdot r_v^2 + \dots + \kappa_n \cdot r_v^n$$

2.3.3 Approximation des Modells

Die Abbildungsfunktion läßt sich auch näherungsweise durch ein Polynom vom Grad n (mit $n > 0$) beschreiben. Je größer der Grad des Polynoms ist, desto kleiner ist der Approximationsfehler. Alle Lichtstrahlen mit dem Winkel θ zur z-Achse haben im Bild den Abstand r vom optischen Zentrum gemäß der folgenden Gleichung:

$$r(\theta) = \kappa_1 \cdot \theta + \kappa_2 \cdot \theta^2 + \dots + \kappa_n \cdot \theta^n$$

Die Aufgabe besteht darin, die Parameter κ_1 bis κ_n zu bestimmen.

2.3.4 Ursachen für Verzerrungen

Das Lochkameramodell ist nur eine Approximation von Kameras mit Linsen. Durch die Verwendung einer Blende vor oder hinter der Linse kommt es auch hier zu radialen Verzerrungen, wenn auch in geringerem Maße. Da die Bildebene planar ist, liegt sie zum Rand hin nicht im Brennpunkt der Linse, wodurch das Bild zum Rand hin unscharf wird. Um eine größere Tiefenschärfe zu erhalten, setzt man eine Blende ein. Die Blende verringert den Durchmesser der Lichtbündels, welches durch die Linse geht und auf die Bildebene trifft. Dadurch verschiebt sich der Mittelpunkt des auf die Bildebene auftreffenden unscharfen Punktes. Dies ergibt eine radiale Verzerrung.

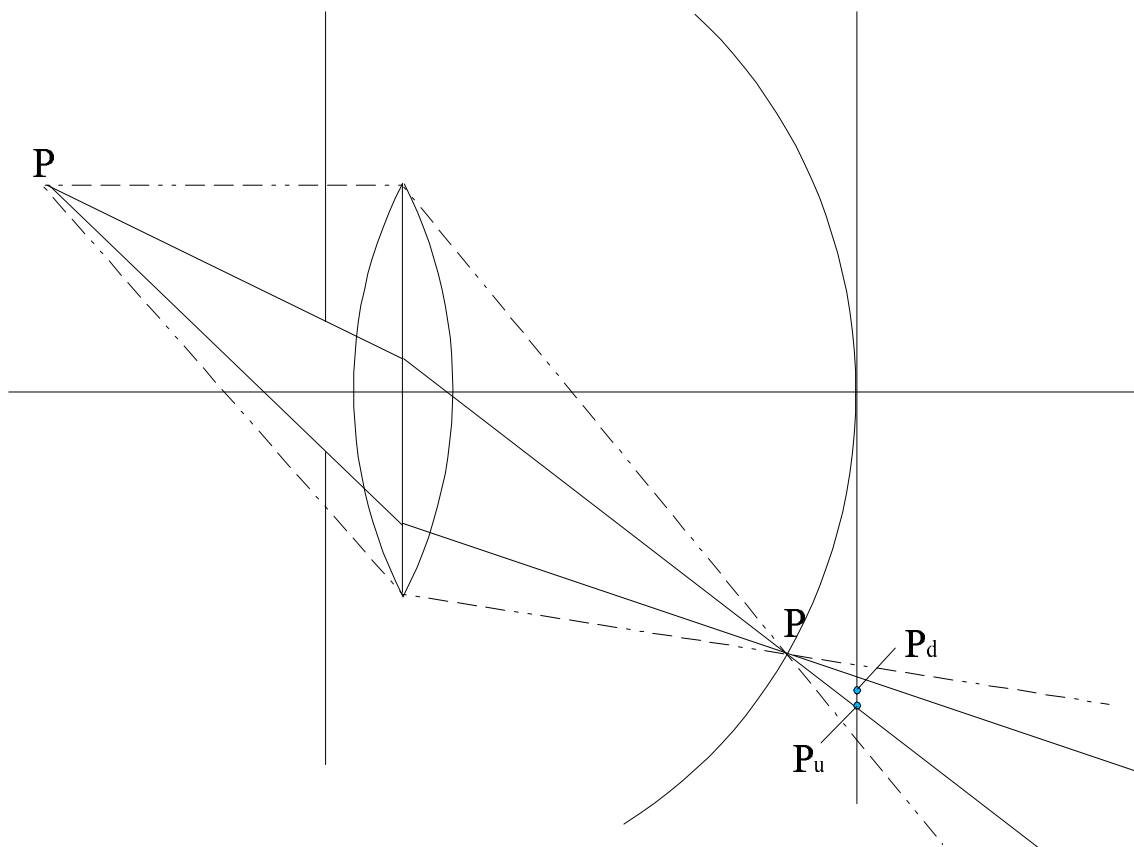


Abbildung 8: Ursache radialer Verzerrungen

2.4 Canny Kantendetektor

Der Canny Kantendetektor arbeitet auf s/w-Bildern. Kanten sind Stellen großer Helligkeitsänderung im Bild, welche mathematisch durch den Gradienten beschrieben werden können. Der Gradient steht senkrecht auf den Kanten und sein Betrag ist ein Maß für die Ausprägtheit einer Kante. Je größer sein Betrag ist, desto stärker ist die Kante ausgeprägt. Um zu entscheiden, ob ein Punkt eine Kante ist, muß man den Betrag seines Gradienten berechnen und mit einem Schwellwert vergleichen. Der Canny-Algorithmus arbeitet sogar mit zwei Schwellwerten. Zuerst wird das Bild mit einem Gaußfilter geglättet, um Rauschen zu unterdrücken. Dann wird für jeden Bildpunkt sein Gradient berechnet. Dies geschieht durch Anwendung des Sobeloperators, der eine Faltungsmaske ist, die Punkt für Punkt über das Bild geschoben wird. Nun wird aus jedem dieser Gradienten sein Betrag berechnet und seine Richtung berechnet. Die Richtung wird benutzt, um an den Kanten entlangzulaufen. Sie hat einen der folgenden Werte: 0, 45, 90 oder 135 Grad. Es seien T_1 und T_2 Schwellwerte mit $T_1 > T_2$ und G der Gradient an einem Punkt. Dann sind alle Punkte mit $|G| > T_1$ Kanten und alle Punkte mit $|G| < T_2$ sind keine Kanten. Alle Bildpunkte mit $T_1 > |G| > T_2$ sind Kanten, falls es einen benachbarten Bildpunkt mit $|G| > T_1$ gibt.

Es sei hier als Beispiel auf Abbildung 9 verwiesen.

3 Berechnung einer Entzerrungsfunktion

3.1 Überblick über das Verfahren

Als Kalibrierungsmuster wird ein schachbrettartiges Muster mit schwarzen und weißen Feldern benutzt. Die Linien, die durch dieses Muster gebildet werden, werden vom Computer erkannt. Ein von der hier verwendeten Kamera mit diesem Kalibrierungsmuster aufgenommenes Bild zeigt Abbildung 9.

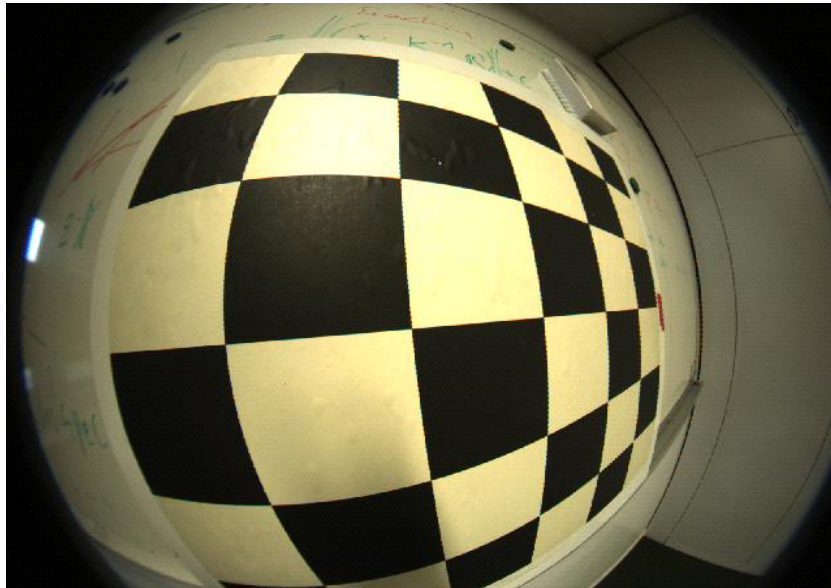


Abbildung 9: Bild des Kalibrierungsmusters, aufgenommen mit einer Fischaugenkamera

Auf dieses Bild wird zunächst der Harris Eckendetektor angewendet. Die von ihm erkannten Ecken entsprechen größtenteils den Berührungspunkten der weißen und schwarzen Felder. Sie werden daher als Startpunkte für die Suche nach den Linien im Bild verwendet. Nun wird der Canny Kantendetektor auf das Bild angewendet. Er liefert ziemlich genau die Linien, die durch die weißen und schwarzen Felder gebildet werden. Das Ergebnis des Kantendetektors angewendet auf das oben gezeigte Bild zeigt die Abbildung 10.

Das vom Kantendetektor gelieferte Bild ist ein binäres Bild. Die weißen Bildpunkte repräsentieren die Kanten. Sie werden im nächsten Arbeitsschritt zu einzelnen Linien segmentiert. Die Möglichkeit des Vorkommens von Lücken in den Kanten wird berücksichtigt. Diese Lücken werden einfach übersprungen, sofern sie eine gewisse Größe nicht überschreiten. Große Lücken führen dazu, daß eine Linie als mehrere Linien erkannt wird. Die Segmentierung wird im nächsten Abschnitt im Detail beschrieben.

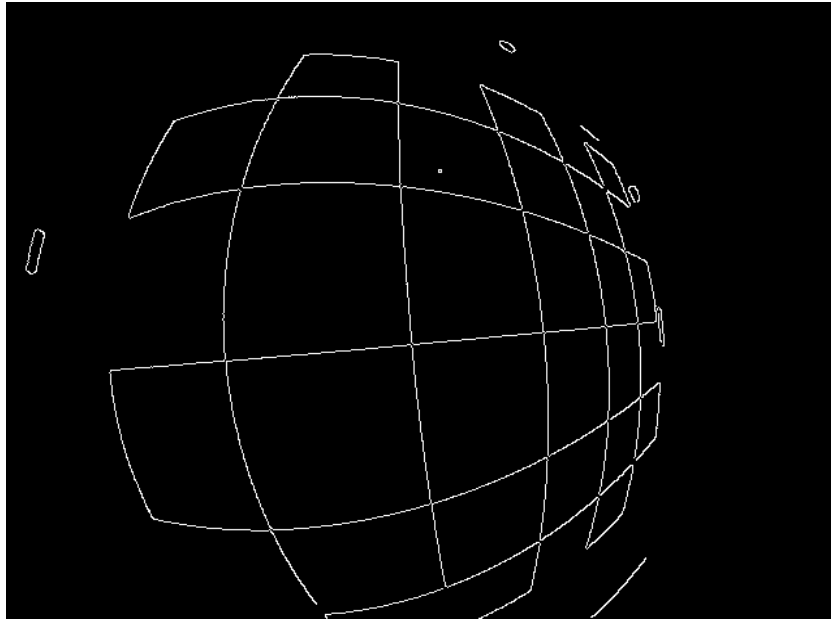


Abbildung 10: Ausgabe des Canny Kantendetektors bei Eingabe des obigen Bildes

Für jede segmentierte Linie wird die zugehörige Regressionsgerade/Regressionsebene lokal berechnet. Diese Regressionsgeraden/-ebenen dienen der Berechnung der Entzerrungsfunktion. Es wird eine Funktion gesucht, die jeden Punkt einer im Bild erkannten Linie auf den zugehörigen Punkt der Regressionsgeraden/-ebene der Linie abbildet. Die gesuchte Funktion wird als Polynom vom Grad n folgender Form dargestellt:

$$r_u = f(r_v) = \kappa_1 \cdot r_v + \kappa_2 \cdot r_v^2 + \dots + \kappa_n \cdot r_v^n$$

Zur Erklärung siehe Kapitel 2.3.2. Diese Gleichung muß für alle Bildpunkte erfüllt sein, wodurch sich ein Gleichungssystem ergibt. Da sich das Gleichungssystem einfacher lösen läßt, wenn das optische Zentrum bekannt ist (es wird dadurch linear), wird dieses bestimmt, bevor das Gleichungssystem gelöst wird. Mit bekanntem Zentrum läßt sich für jeden Bildpunkt sein Abstand zu diesem berechnen. Für jeden Punkt einer Linie erhält man somit zwei Abstände zum Zentrum: r_u (unverzerrt) und r_v (verzerrt). Die Werte für κ_1 bis κ_n werden mit Hilfe der Methode der kleinsten Quadrate bestimmt. Für Details siehe Kapitel 3.4.

Abbildung 11 zeigt die für das in Abbildung 9 gezeigte Bild erkannten Linien mit ihren Regressionsgeraden.

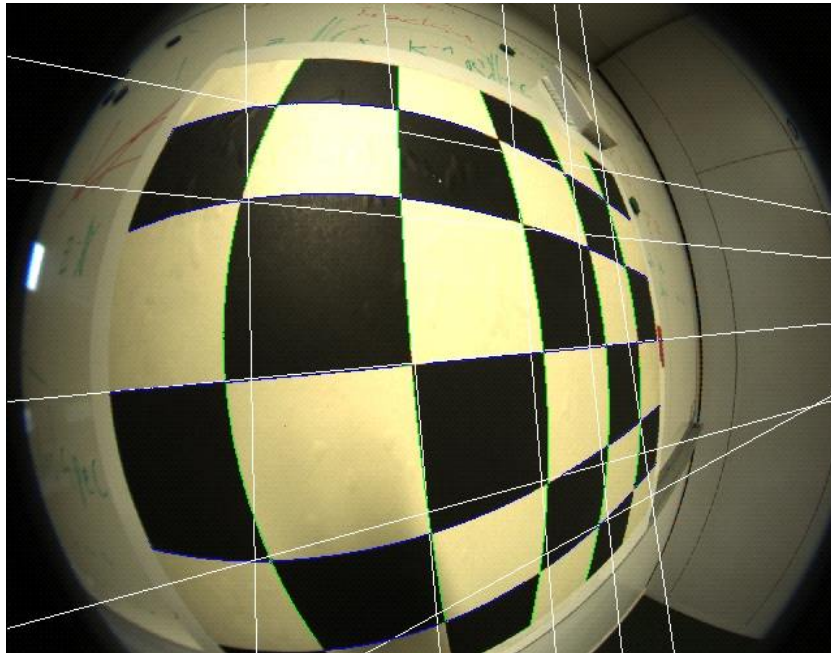


Abbildung 11: Erkannte Linien (grün und blau markiert) mit zugehörigen Regressionsgeraden (weiß).

3.2 Bestimmung der Linien des Kalibrierungsmusters

Die weißen und schwarzen Felder des Kalibrierungsmusters bilden Kanten, die vom Canny Kantendetektor gut erkannt werden. In einem unverzerrtem Bild entsprächen diese Kanten geraden Linien. Durch die Verzerrung, die die Kamera hervorruft, werden diese Kanten krumm. Da wir wissen, daß diese Kanten gerade sein sollen, können wir mit ihnen die Entzerrungsfunktion der Kamera berechnen. Im vom Kantendetektor gelieferten Bild sind alle Linien miteinander verbunden, so daß sie erst noch voneinander getrennt werden müssen, um für jede Linie eine Regressionsgerade berechnen zu können. Ausgehend von den vom Eckendetektor gelieferten Ecken wird nach rechts, links, oben und unten nach Linien im vom Kantendetektor berechnetem Bild gesucht.

Für jede erkannte Ecke wird folgendes gemacht:

1. Suche alle weißen Punkte in unmittelbarer Umgebung rechts von der Ecke, füge diese Punkte in eine Schlange ein und berechne für diese Punkte eine Regressionsgerade. (Ein Suchbereich von 10 x 20 Bildpunkten hat sich als gut erwiesen. Der Suchbereich ist asymmetrisch, da im Programm zwischen horizontalen und vertikalen Linien unterschieden wird. Diese Unterscheidung ist notwendig, da die Geraden in der Koordinatendarstellung implementiert sind.)
2. Gehe eine Spalte weiter nach rechts im Bild und suche nach dem weißen Punkt, der den kleinsten Abstand zur Regressionsgeraden hat und füge diesen in die Schlange ein. Wurde kein solcher Punkt gefunden, wiederhole Schritt 2 bis zu n mal. Findet sich auch beim n-ten Mal kein solcher Punkt, fahre mit Schritt 5 fort.
3. Entferne den letzten Punkt aus der Schlange und berechne für die nun sich in der Schlange befindlichen Punkte eine Regressionsgerade.
4. Fahre mit Schritt 2 fort.
5. Suche alle weißen Punkte in unmittelbarer Umgebung links von der Ecke, füge diese Punkte in eine Schlange ein und berechne für diese Punkte eine Regressionsgerade.

Dies wird analog nach links, oben und unten gemacht. Nach diesen Schritten ist eine

Linie segmentiert. Für die restlichen Ecken wird dasselbe gemacht. Danach sind alle Linien im Bild segmentiert. Lücken in dem vom Kantendetektor geliierten Linien werden nicht aufgefüllt. In Schritt 2 wird berücksichtigt, daß es Lücken geben kann. Diese werden jedoch einfach übersprungen, sofern sie eine Größe von n Punkten nicht übersteigen. Durch solche Lücken kann es passieren, daß eine Linie bei der Segmentierung in mehrere Linien aufgeteilt wird.

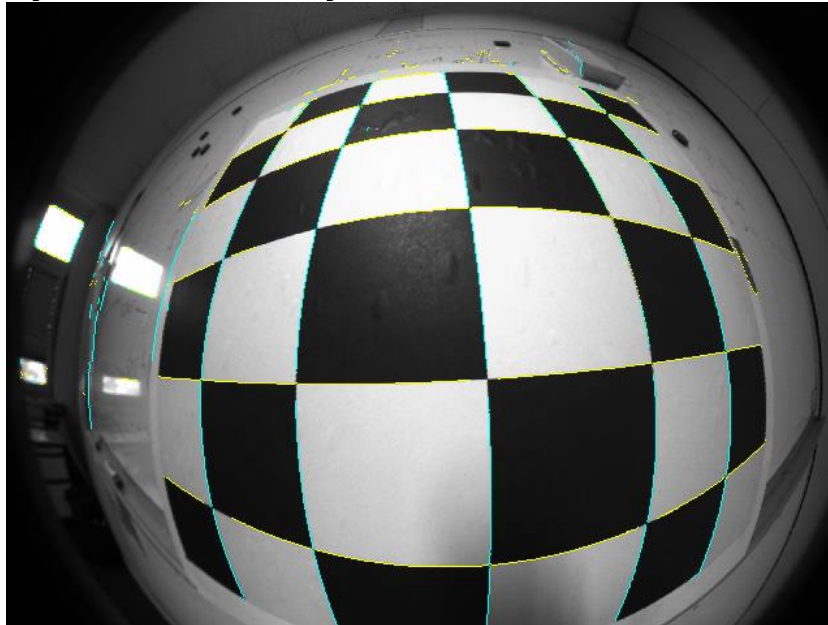


Abbildung 12: Erkannte Linien, gelb und blau markiert.

3.3 Bestimmung des optischen Zentrums

Das optische Zentrum einer Kamera ist nicht notwendigerweise in der Bildmitte. Ist das optische Zentrum bekannt, dann ist zur Bestimmung der Verzerrungsfunktion nur die Minimierung eines linearen Gleichungssystems notwendig. Deswegen wird bei dem hier vorgestellten Verfahren das optische Zentrum bestimmt. Das optische Zentrum zeichnet sich dadurch aus, daß gerade Linien im dreidimensionalen Raum, die es im Kamerabild durchkreuzen, dort ebenfalls gerade sind, also durch die Verzerrung nicht gekrümmt werden. Für jede im Bild erkannte Linie wird im Vergleich mit ihrer Regressionsgeraden die durchschnittliche Entfernung aller ihrer Punkte zur Regressionsgeraden bestimmt. Für Linien, die durch das optische Zentrum gehen, ist diese durchschnittliche Entfernung nahezu gleich Null. Um das optische Zentrum möglichst genau bestimmen zu können, ist eine Sequenz von Bildern notwendig, bei denen einmal eine horizontale und einmal eine vertikale Linie des Kalibrierungsmusters das optische Zentrum im Bild kreuzen. Mit horizontalen Linien wird die y -Koordinate, mit vertikalen Linien die x -Koordinate des optischen Zentrums bestimmt. Das beste Ergebnis aus einer Bildsequenz wird gespeichert und für die weiteren Berechnungen benutzt.

3.4 Bestimmung der radialsymmetrischen Entzerrungsfunktion

Dieses Verfahren ist für Fischaugenkameras optimiert und geht daher von einer radialsymmetrischen Verzerrung des Bildes aus. Es wird ausgenutzt, daß gerade Linien im dreidimensionalen Raum bei der Projektion auf die Bildebene einer Lochkamera unabhängig von der Position und der Lage der Kamera gerade bleiben. Die Entzerrungsfunktion ist ein Polynom vom Grad n und hat die Form

$$r_u = f(r_v) = \kappa_1 \cdot r_v + \kappa_2 \cdot r_v^2 + \dots + \kappa_n \cdot r_v^n .$$

Das in dieser Studienarbeit entwickelte Programm arbeitet in einem von 3 Modi:

Modus 1 ist für Bilder nach dem Lochkameramodell. Die Abstände entsprechen der euklidischen Norm. Zu jeder im Bild erkannten Linie wird eine Ausgleichsgerade berechnet.

Modus 2 ist für Bilder nach dem Kugelkameramodell. Jedes Bild wird zunächst in die Sicht entsprechend dem Lochkameramodell umgerechnet und es wird dann wie im ersten Modus weiter verfahren. Aufgrund der abnehmenden Auflösung zum Bildrand hin ist dies jedoch problematisch.

Modus 3 ist auch für Bilder nach dem Kugelkameramodell. Das Programm arbeitet intern mit einer Theta-Phi-Darstellung des Kamerabildes. Der Abstand eines Bildpunktes zum optischen Zentrum im Bild entspricht dem Winkel θ (in Kugelkoordinaten). Alle Bildpunkte werden als Punkte auf einer Kugel (mit dem Radius 1) angesehen. Der Abstand zwischen zwei Punkten entspricht der Länge des Kreisbogens, der sich durch den Winkel zwischen den beiden Punkten ergibt (Winkel in Radiant). Anstelle von Regressionsgeraden werden Regressionsebenen berechnet. Der Abstand zwischen einem Punkt und einer Ebene wird ebenfalls als Winkel/Kreisbogen gemessen.

Gegeben sind die Linien des Kalibrierungsmusters sowie das optische Zentrum. Für jede Linie des Kalibrierungsmusters wird eine Ausgleichsgerade bzw. eine Ausgleichsebene berechnet. Die hier bestimmte Funktion ordnet jedem Punkt der Linie möglichst einem Punkt auf der Ausgleichsgeraden/Ausgleichsebene zu. Wird die Funktion über alle Linien des Kalibrierungsmusters minimiert, so liefert sie auch für alle übrigen Punkte des Bildes brauchbare Ergebnisse.

Es bezeichne r_v den Abstand eines Punktes einer Linie des Kalibrierungsmusters vom optischen Zentrum und r_u den Abstand des zugehörigen Punktes auf der Ausgleichsgeraden/Ausgleichsebene vom optischen Zentrum. Dann erhält man ein lineares Gleichungssystem der Form

$$\begin{bmatrix} r_{v,1} & r_{v,1}^2 & r_{v,1}^3 & \dots & r_{v,1}^n \\ r_{v,2} & r_{v,2}^2 & r_{v,2}^3 & \dots & r_{v,2}^n \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ r_{v,m} & r_{v,m}^2 & r_{v,m}^3 & \dots & r_{v,m}^n \end{bmatrix} \cdot \begin{bmatrix} \kappa_1 \\ \kappa_2 \\ \vdots \\ \kappa_n \end{bmatrix} = \begin{bmatrix} r_{u,1} \\ r_{u,2} \\ \vdots \\ r_{u,m} \end{bmatrix}$$

mit $n = \text{Grad des Polynoms}$ und $m = \text{Anzahl der Punkte}$.

Es werden bessere Ergebnisse erzielt, wenn die r_u so skaliert werden, daß die Ausgleichsgerade zu einer Tangente der detektierten Linie wird. Es werden die r_v und r_u gesucht, für die gilt: $(r_u - r_v) \rightarrow \min$. Diese werden im folgenden mit $r_{u \min}$ und $r_{v \min}$ bezeichnet. Die Skalierung von r_u erfolgt dann durch: $r_u =: r_u * r_{u \min} / r_{v \min}$. Die Werte für κ_1 bis κ_n werden erst nach der Skalierung der r_u s nach dem

Verfahren der kleinsten Quadrate berechnet. Durch diese Werte ist die Entzerrungsfunktion eindeutig bestimmt.

Eine andere Möglichkeit, dieses Verfahren noch zu verbessern, ist, sich iterativ einem optimalen Ergebnis zu nähern. Dazu werden die erkannten Linien mit der im ersten Durchgang bestimmten Entzerrungsfunktion entzerrt und für diese entzerrten Linien werden Ausgleichsgeraden/-ebenen berechnet. Im nächsten Iterationsschritt werden die neu berechneten Ausgleichsgeraden/-ebenen anstelle der Ausgleichsgeraden/-ebenen der erkannten Linien zur Berechnung einer Entzerrungsfunktion benutzt. Die r_v s bleiben also in allen Iterationsschritten gleich, während die r_u s in jedem Schritt aus den neu berechneten Ausgleichsgeraden/-ebenen berechnet werden. Bei diesem iterativen Verfahren wird die oben beschriebene Skalierung nicht durchgeführt. Sie ergibt sich automatisch durch die Iteration. Die Krümmung der Linien nach der Entzerrung ist ein Maß dafür, wie gut die Approximation der Entzerrungsfunktion ist. Das Programm prüft nach jedem Iterationsschritt an Hand dieses Kriteriums, ob die Iteration abgebrochen werden soll. Die maximale Anzahl an Iterationsschritten ist begrenzt für den Fall, daß die Iteration nicht konvergiert.

Da die Entzerrungsfunktion streng monoton steigend und daher bijektiv ist, läßt sie sich invertieren, wodurch man die Entzerrungsfunktion erhält.

4 Ergebnisse und Bewertung

4.1 Ergebnisse

Die entzerrten Bilder lassen sich auf zwei verschiedene Weisen berechnen. Das Einfachste ist es, für jeden Bildpunkt im Originalbild durch Anwendung der Entzerrungsfunktion die neuen Koordinaten für den Punkt zu berechnen und ihn an die errechnete Stelle im Zielbild einzusetzen. Diese Methode wird *forward mapping* genannt. Ihr Nachteil ist, daß im Zielbild schwarze Linien auftauchen, weil nicht jeder Punkt im Originalbild einem Punkt im Zielbild zugeordnet wird. Man kann dieses Problem lösen, indem man umgekehrt vorgeht, also für jeden Punkt im Zielbild den zugehörigen Punkt im Originalbild berechnet und diesen ins Zielbild einsetzt. Dazu muß man die Entzerrungsfunktion invertieren, da man eine Verzerrungsfunktion braucht. Die errechneten Koordinaten im Originalbild sind im allgemeinen keine ganzen Zahlen, so daß man über die Farbwerte der benachbarten Punkte interpoliert, um einen Wert für den Punkt im Zielbild zu erhalten. Man nennt dieses Verfahren *backward mapping*.

Das Programm wurde zunächst mit einer herkömmlichen Kamera getestet, da die Verzerrungen damit geringer sind. Abbildung 13 zeigt das aufgenommene Bild.

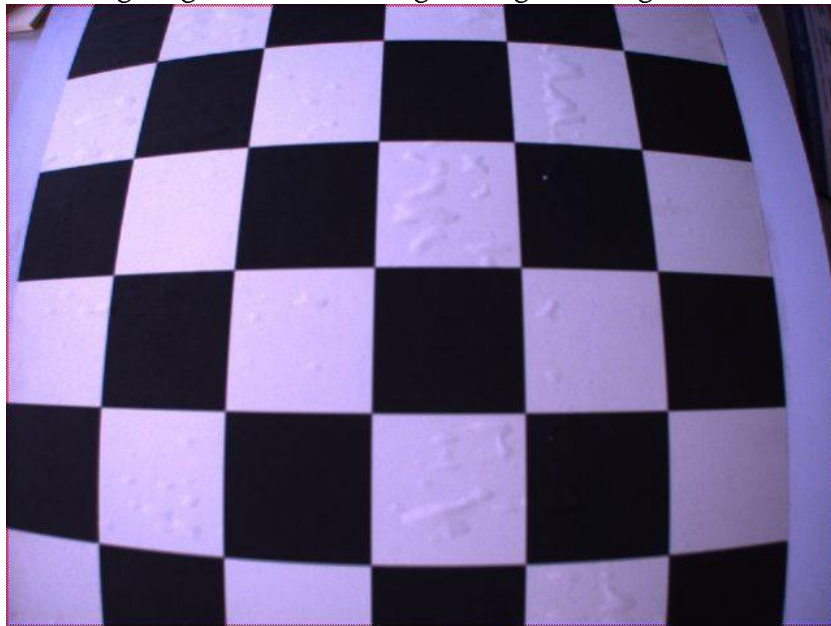


Abbildung 13: Mit einer dem Lochkameranmodell entsprechenden Kamera aufgenommenes Bild. Die Verzerrungen sind besonders am Rand stark ausgeprägt.

Ohne den iterativen Ansatz und ohne die Skalierung der r_u s berechnet das Programm das folgende Bild:

Wird die Skalierung der r_u s durchgeführt, do erhält man das folgende Bild:

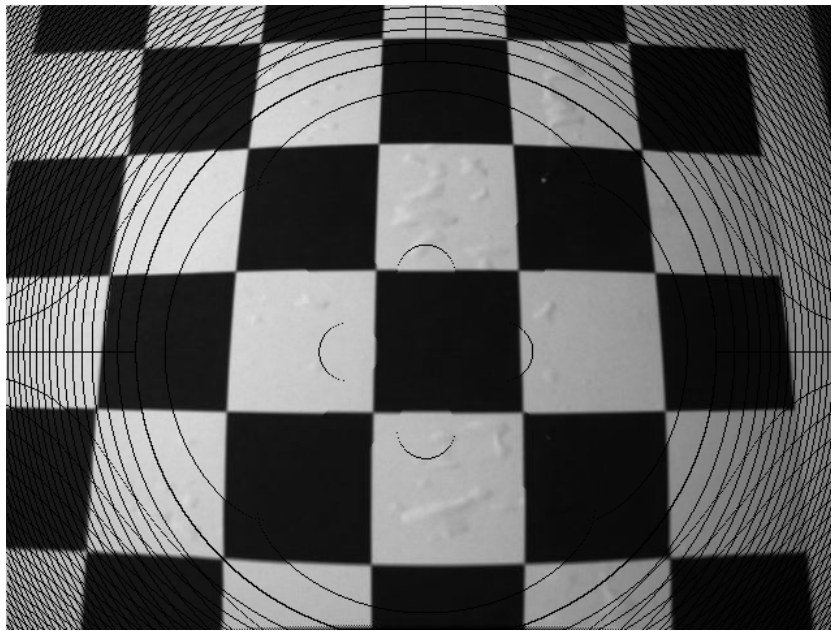


Abbildung 14: Vom Programm entzerrtes Bild.

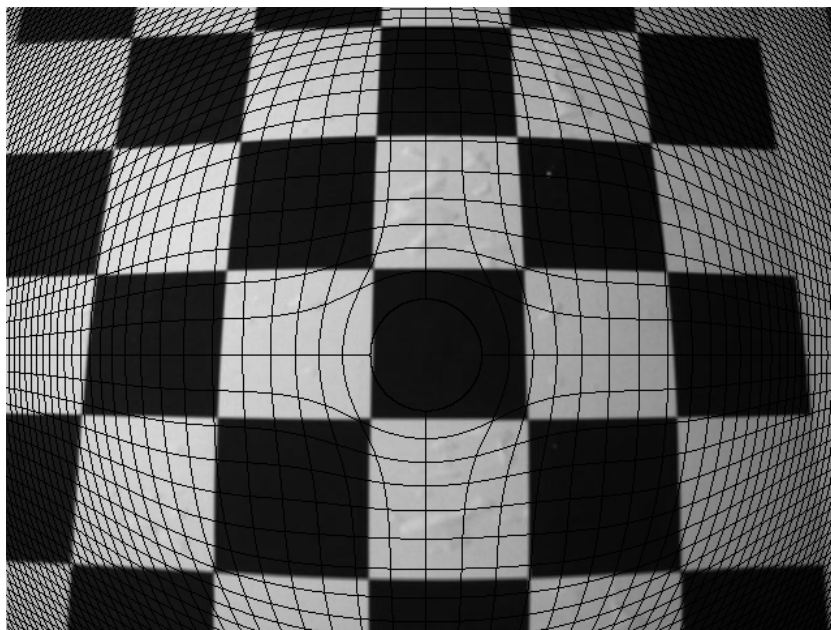


Abbildung 15: Vom Programm entzerrtes Bild mit Skalierung der Regressionsgeraden

Dieses Ergebnis ist deutlich besser als das ohne Skalierung. Führt man die Bestimmung der Entzerrungsfunktion iterativ durch, so erhält man bereits nach wenigen Iterationsschritten ziemlich gute Ergebnisse, wie das folgende Bild zeigt. Es wurde mit 5 Iterationen erstellt. Zum Vergleich folgt das selbe Bild nach dem *backward mapping*-Verfahren berechnet.

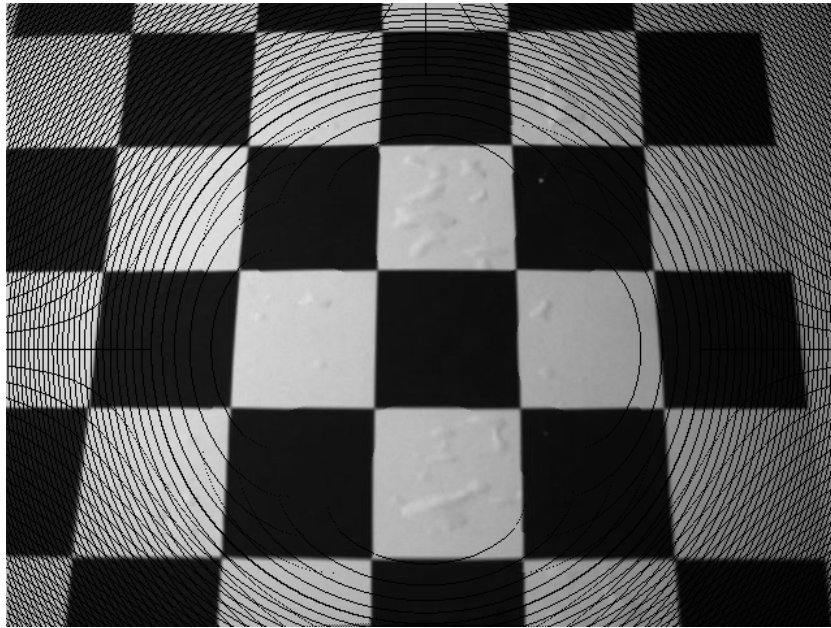


Abbildung 16: Vom Programm nach dem iterativen Verfahren entzerrtes Bild. (forward mapping)

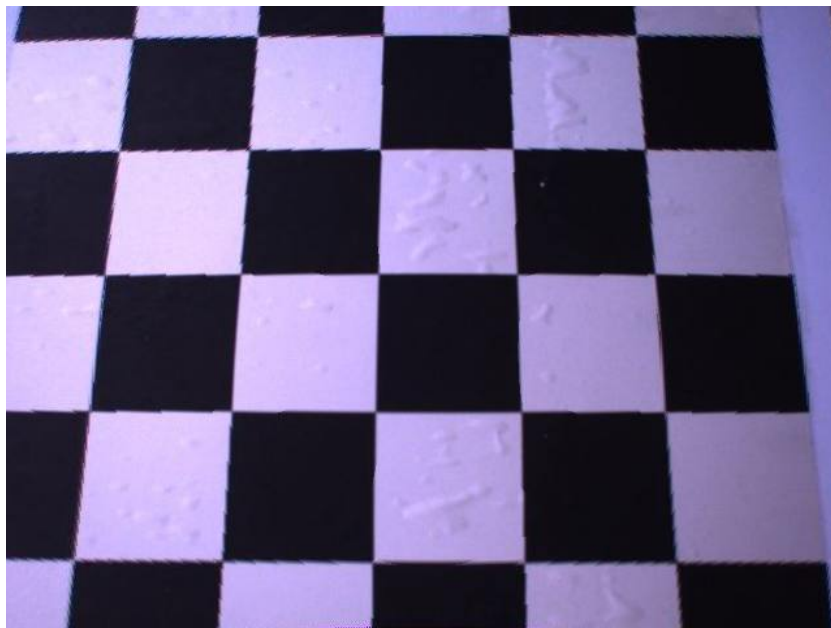


Abbildung 17: Vom Programm nach dem iterativen Verfahren entzerrtes Bild. (backward mapping)

Die errechnete Entzerrungsfunktion ist:

$$r_u = 0.8880950153 \cdot r_v + 0.001357729048 \cdot r_v^2 - 1.069665102e-05 \cdot r_v^3 + 4.182428797e-08 \cdot r_v^4 - 4.609097348e-11 \cdot r_v^5 \quad (f1)$$

Die folgende Abbildung zeigt eine graphische Darstellung dieser Funktion.

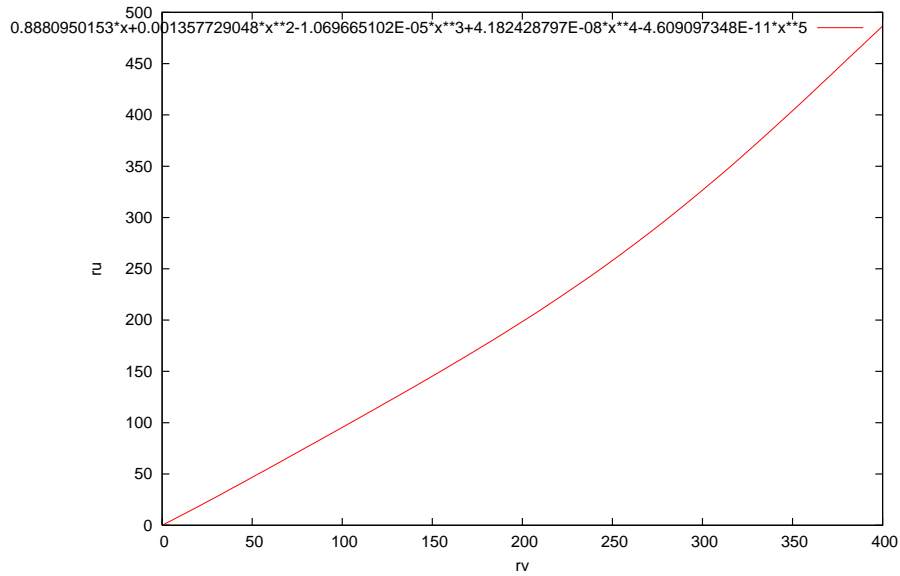


Abbildung 18: Graphische Darstellung der Funktion f_1

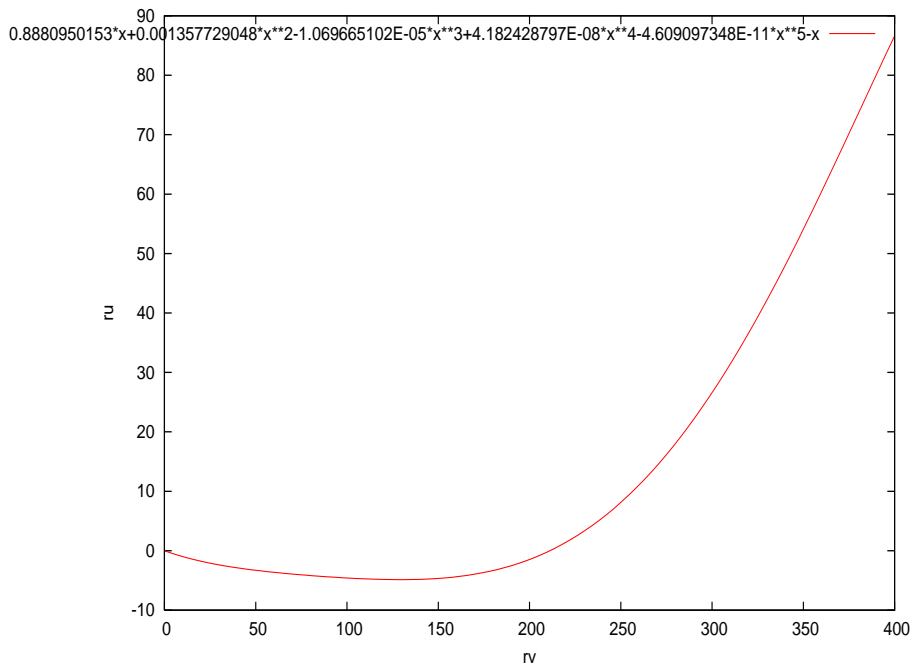
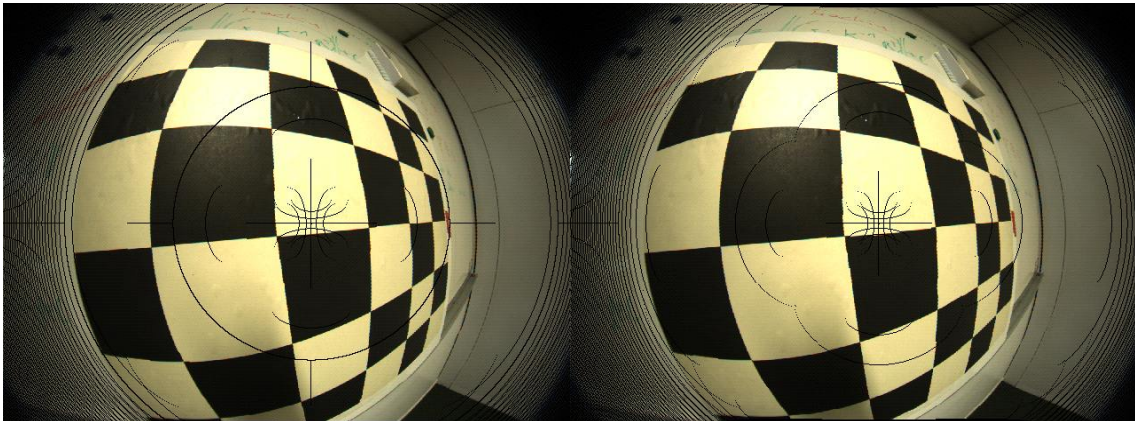


Abbildung 19: Graphische Darstellung der Abweichung $(f_1 - r)$

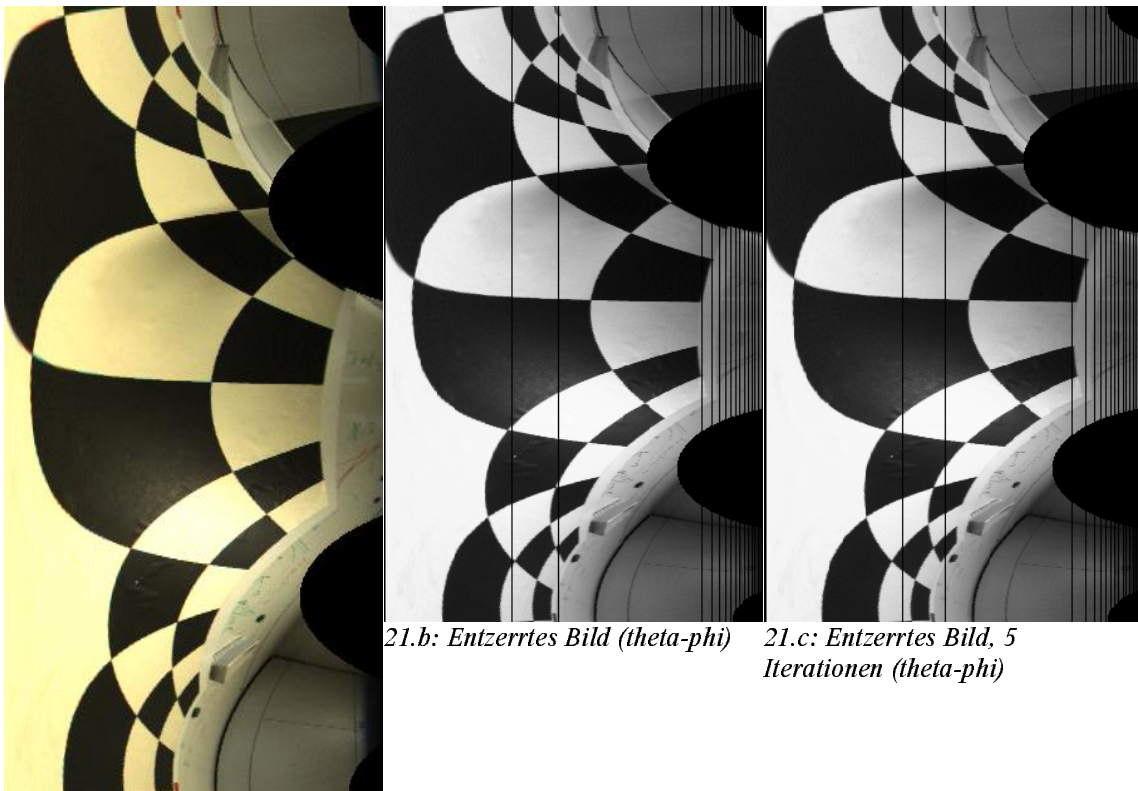
Die folgenden Abbildungen zeigen die berechneten entzerrten Bilder für ein mit einer Fischaugenkamera aufgenommenes Bild.



20.a: Entzerrtes Bild

20.b: Entzerrtes Bild, 5 Iterationen

Abbildung 20: Entzerrtes Bild mit und ohne iterativem Verfahren. Diese Bilder basieren auf dem in Abbildung 9 gezeigten Bild.

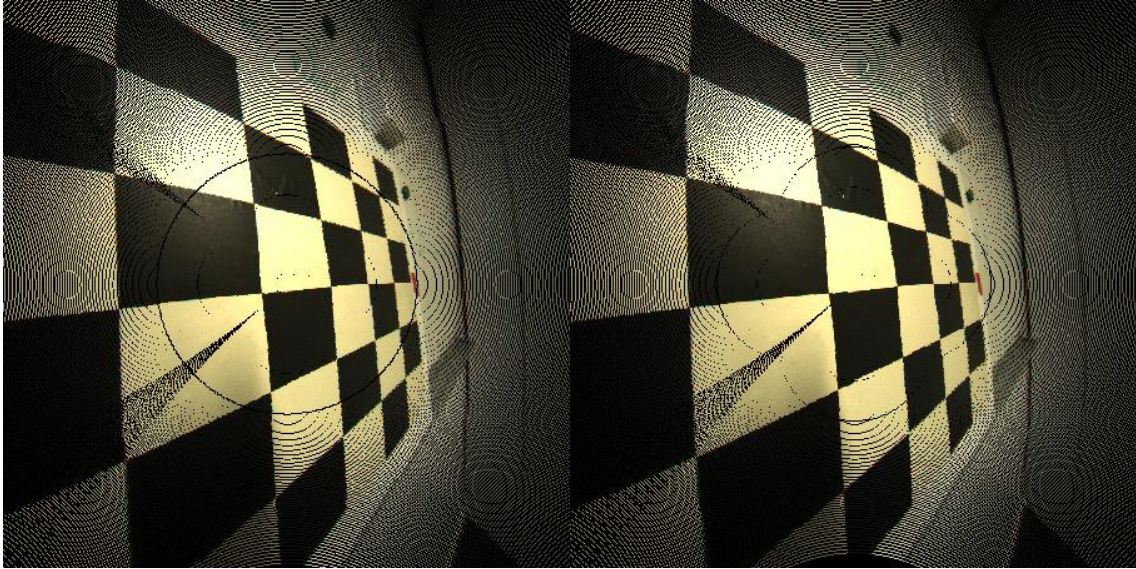


21.a: Originalbild (θ - ϕ)

21.b: Entzerrtes Bild (θ - ϕ)

21.c: Entzerrtes Bild, 5 Iterationen (θ - ϕ)

Abbildung 21: Entzerrtes Bild mit und ohne iterativem Verfahren (θ - ϕ). Diese Bilder basieren auf dem in Abbildung 9 gezeigten Bild.



22.a: Entzerrtes Bild aus der Sicht einer Lochkamera

22.b: Entzerrtes Bild aus der Sicht einer Lochkamera , 5 Iterationen

Abbildung 22: Entzerrtes Bild aus der Sicht einer Lochkamera mit und ohne iterativem Verfahren

Die errechnete Entzerrungsfunktion (ohne Iteration) ist:

$$\theta_u = 1.243108261 \cdot \theta_v - 1.377258184 \cdot \theta_v^2 + 2.638620731 \cdot \theta_v^3 - 2.085702893 \cdot \theta_v^4 + 0.5825628844 \cdot \theta_v^5 \quad (f2)$$

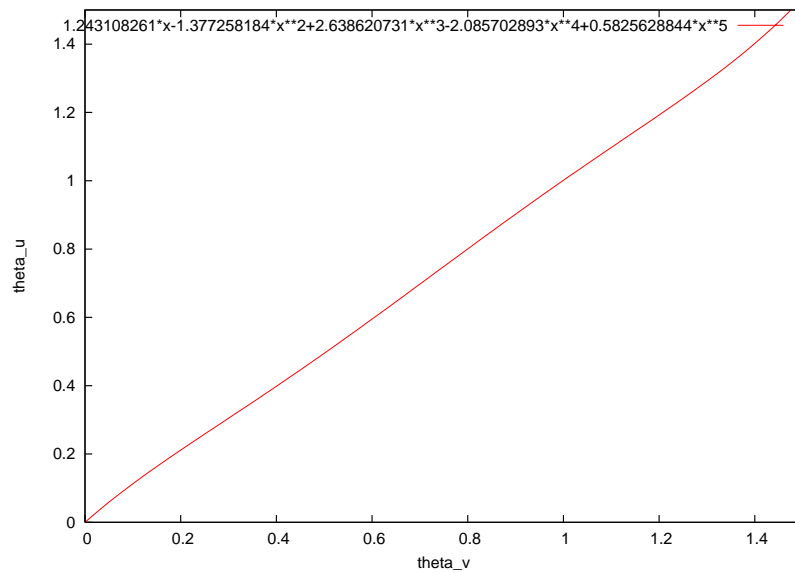


Abbildung 23: Graphische Darstellung der Funktion f2.

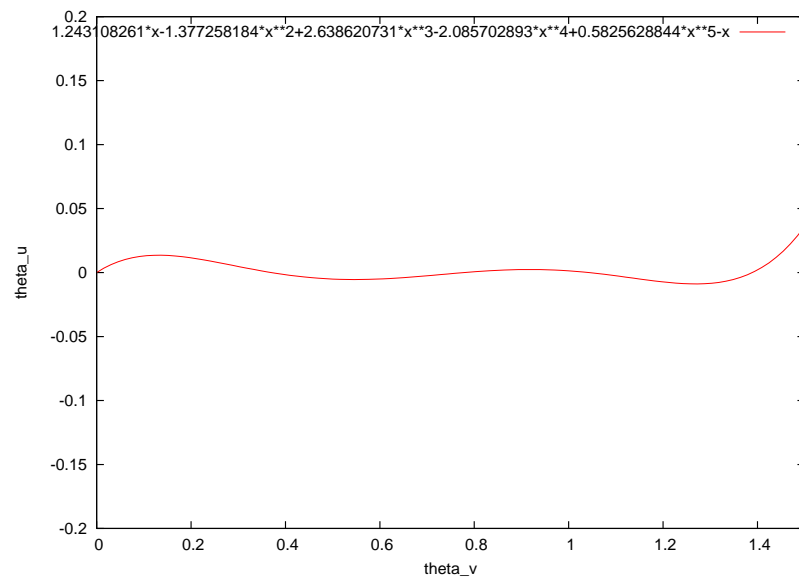


Abbildung 24: Graphische Darstellung der Abweichung ($f_2 - \theta$)

Nach fünfacher Iteration errechnet das Programm folgende Entzerrungsfunktion:

$$\theta_u = 1.270835985 \cdot \theta_v - 1.643125213 \cdot \theta_v^2 + 3.148384759 \cdot \theta_v^3 - 2.467542738 \cdot \theta_v^4 + 0.6812626772 \cdot \theta_v^5 \quad (f2Iter)$$

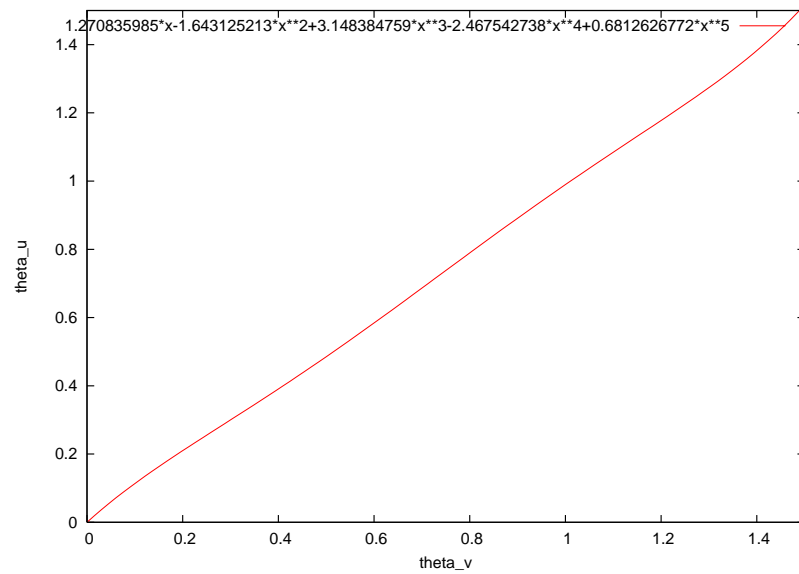


Abbildung 25: Graphische Darstellung der Funktion f_2Iter .

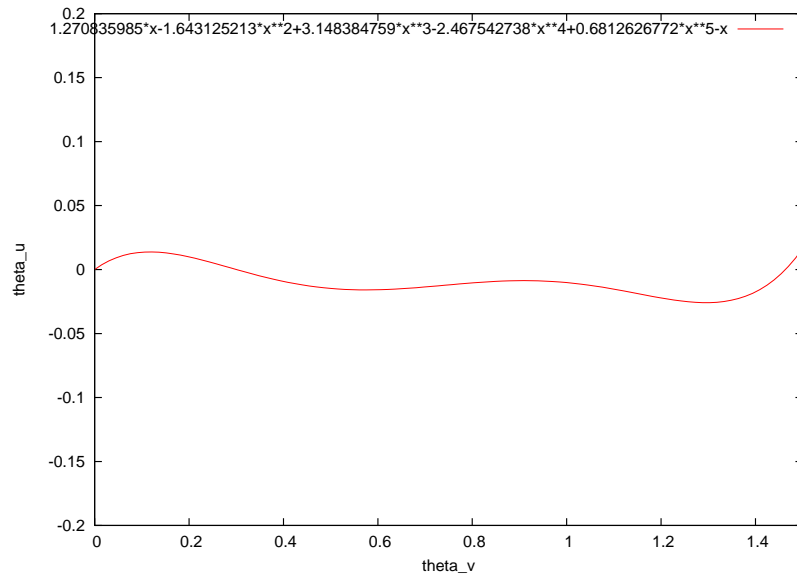


Abbildung 26: Graphische Darstellung der Abweichung ($f2Iter-\theta$).

Einem weiteren Verfahren für Bilder von Fischaugenkameras entsprechend, welches auch implementiert wurde, wird das Bild zunächst in die Sicht einer Lochkamera umgerechnet. Dies wurde für das in Abbildung 9 gezeigte Bild durchgeführt.

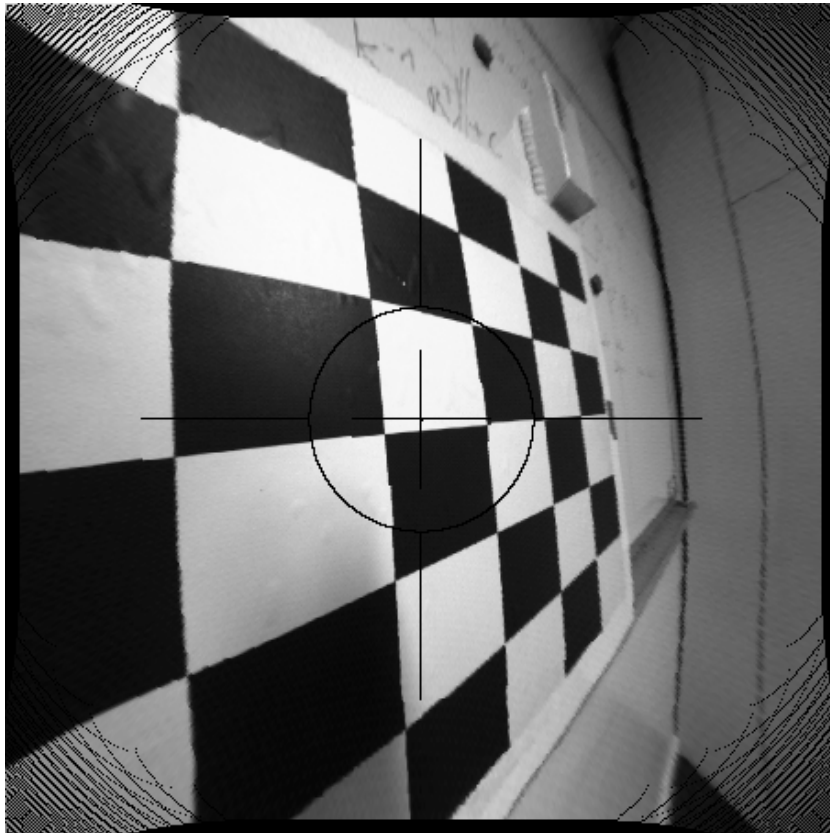


Abbildung 27: Entzerrtes Bild, basiert auf Bild aus Abb. 9, welches zunächst in die Sicht einer Lochkamera umgerechnet wurde.

Das Programm berechnet für das in Abb. 9 gezeigte Bild nach diesem Verfahren die folgende Entzerrungsfunktion:

$$r_u = 1.03287209 \cdot r_v - 0.001713988501 \cdot r_v^2 + 2.854197319 \cdot 10^{-5} \cdot r_v^3 - 1.718628461 \cdot 10^{-7} \cdot r_v^4 + 3.227987626 \cdot 10^{-10} \cdot r_v^5 \quad (f3)$$

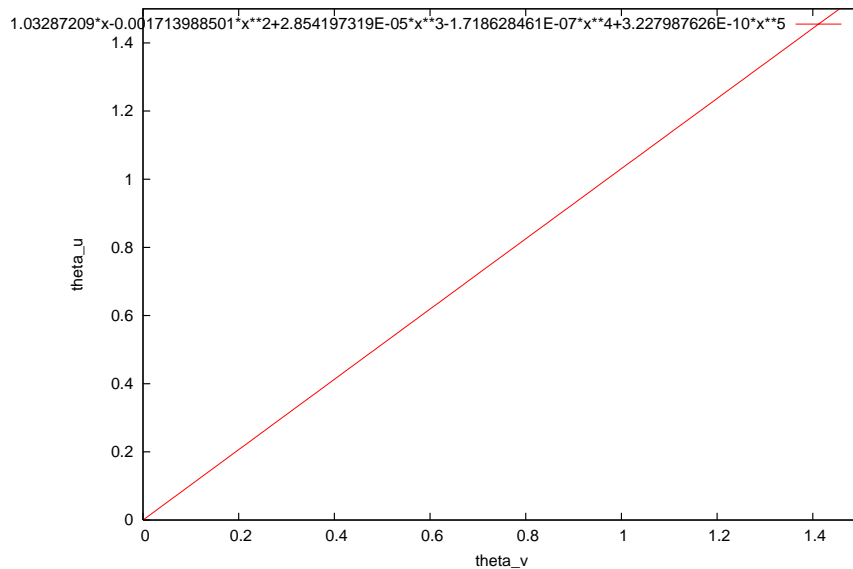


Abbildung 28: Graphische Darstellung der Funktion $f3$.

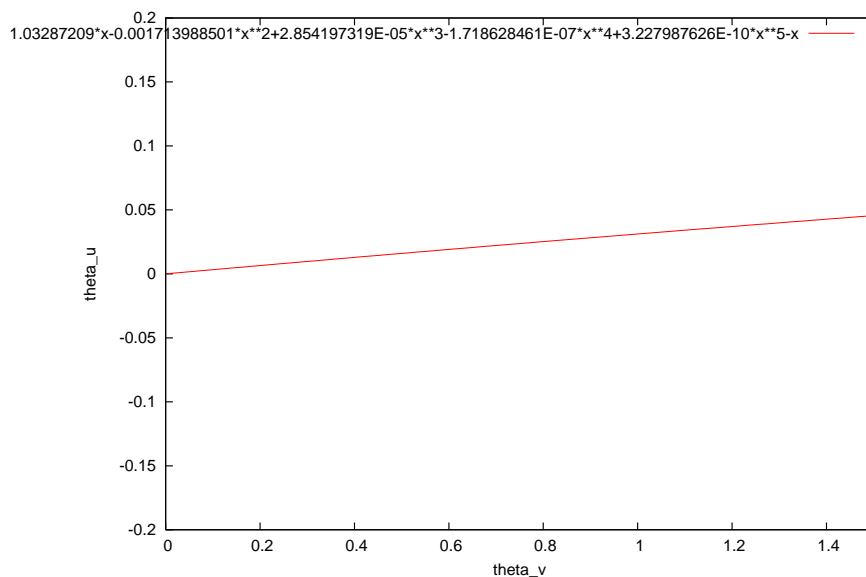


Abbildung 29: Graphische Darstellung der Abweichung $(f3-r)$.

Aufgrund der chromatischen Abberation ist es sinnvoll, die Entzerrungsfunktion für die Farben Rot, Grün und Blau getrennt zu berechnen. Versuche haben gezeigt, daß die berechneten Funktionen für die drei Grundfarben leicht voneinander abweichen. Die chromatische Abberation ist nachweislich ein nicht zu unterschätzender Faktor bei der Berechnung der Entzerrungsfunktion. Für das in Abbildung 12 gezeigte Bild wurden Entzerrungsfunktionen für den Rot-, Grün- und Blauanteil im Bild berechnet. Das Programm verwendet dabei nur die Bilddaten jeweils eines Kanals. Berechnete Entzerrungsfunktion für den roten Kanal:

$$\theta_u = 0.9506801447 * \theta_v - 0.09589612507 * \theta_v^2 + 0.7770242663 * \theta_v^3 - 0.9914805741 * \theta_v^4 + 0.3672187859 * \theta_v^5 \quad (\text{f4R})$$

Berechnete Entzerrungsfunktion für den grünen Kanal:

$$\theta_u = 0.9085711214 * \theta_v + 0.2824791412 * \theta_v^2 - 0.2219253491 * \theta_v^3 + 0.02700326513 * \theta_v^4 + 0.01362425064 * \theta_v^5 \quad (\text{f4G})$$

Berechnete Entzerrungsfunktion für den blauen Kanal:

$$\theta_u = 1.0145142 * \theta_v - 0.5953126483 * \theta_v^2 + 1.975118512 * \theta_v^3 - 2.123249681 * \theta_v^4 + 0.7373518404 * \theta_v^5 \quad (\text{f4B})$$

Berechnete Entzerrungsfunktion für Grauwertbild:

$$\theta_u = 0.9047364236 * \theta_v + 0.2022331597 * \theta_v^2 + 0.1503393387 * \theta_v^3 - 0.4638199966 * \theta_v^4 + 0.2130220357 * \theta_v^5 \quad (\text{f4s/w})$$

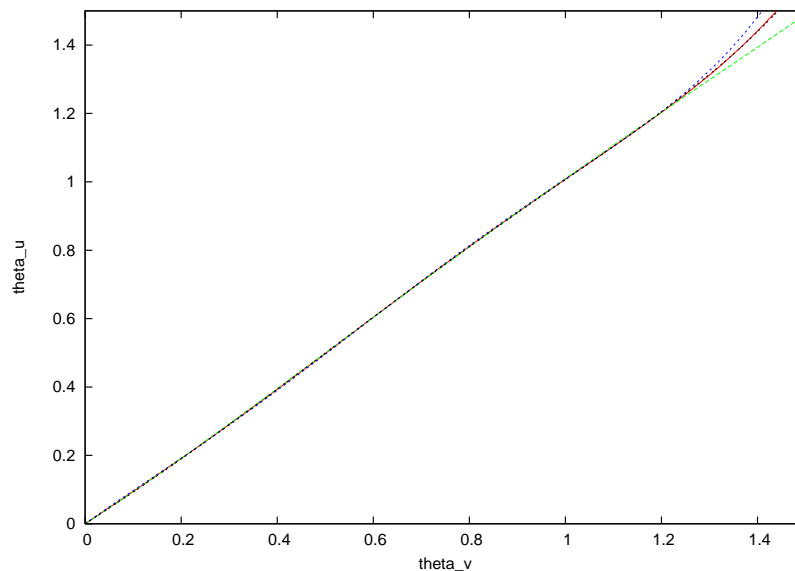


Abbildung 30: Graphische Darstellung der Funktionen f4R, f4G, f4B und f4s/w

Um einen möglichst großen Wertebereich der Abstände r_v abzudecken, empfiehlt es sich, die im Bild erkannten Linien über eine Sequenz von Bildern zu sammeln und die Berechnung der Entzerrungsfunktion über alle in einer Sequenz gefundenen Linien durchzuführen. Über die in Abbildung 31 gezeigte Bildsequenz wurde eine Entzerrungsfunktion berechnet.

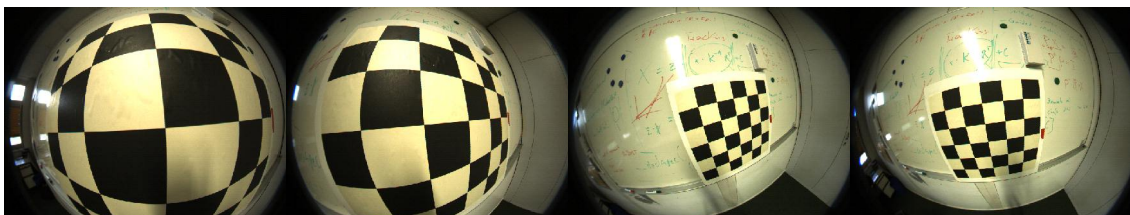


Abbildung 31: Bildsequenz

Als Ergebnis erhält man die folgende Funktion:

$$\theta_u = 1.111653282 * \theta_v - 0.6769587151 * \theta_v^2 + 1.378199928 * \theta_v^3 - 1.113878154 * \theta_v^4 + 0.3099286944 * \theta_v^5 \quad (f5)$$

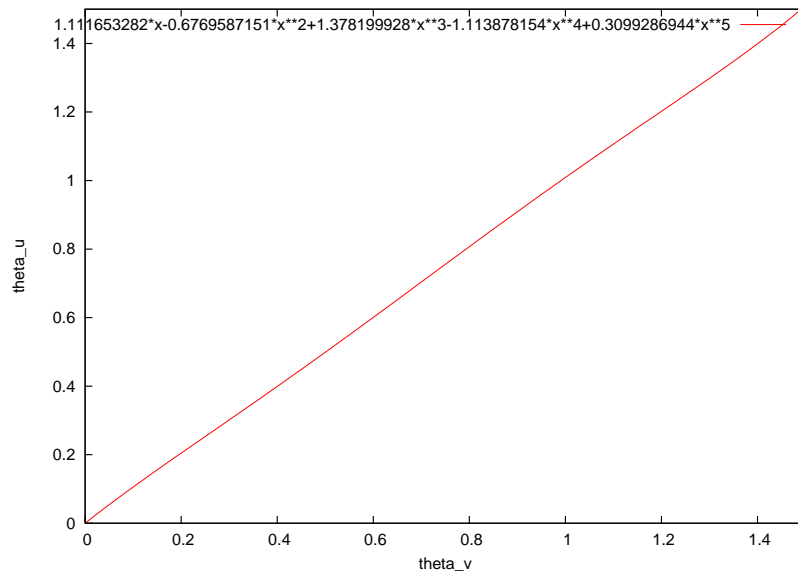


Abbildung 32: Graphische Darstellung der Funktion f_5

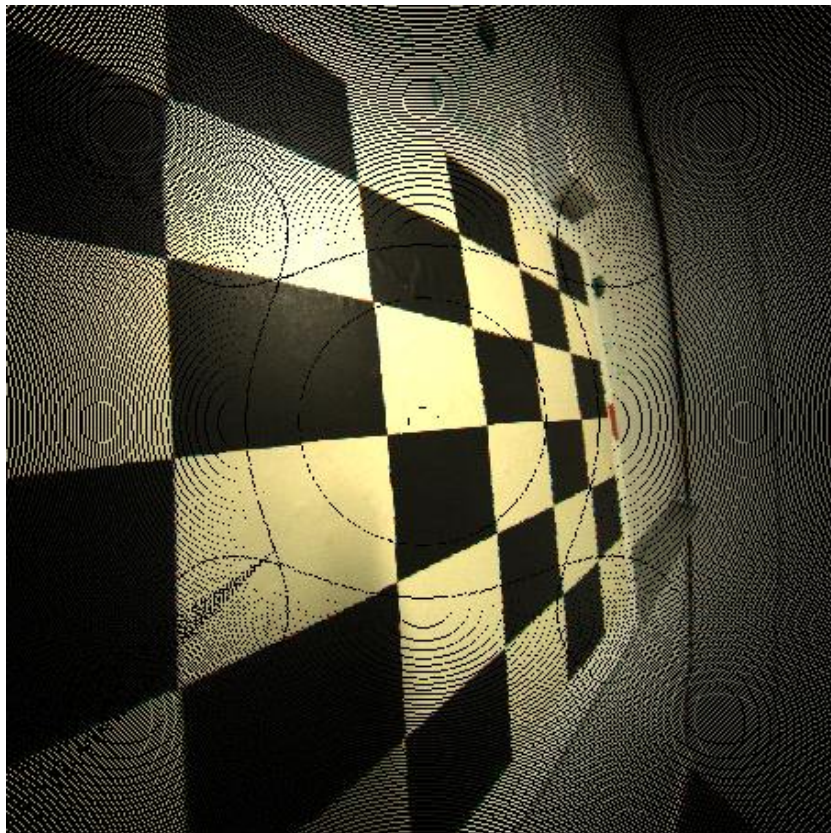


Abbildung 33: Entzerrtes Bild aus der Sicht einer Lochkamera (basiert auf Bildsequenz aus Abb.31)

4.2 Bewertung

Es hat sich gezeigt, daß das Kalibrierungsmuster möglichst das ganze Kamerabild füllen muß, damit es für einen möglichst großen Wertebereich gemessene Abstände r_v bzw.

θ_v gibt. Gibt man dem Programm z. B. nur eine Linie als Eingabe, so erhält man eine Funktion, die zwar für alle Punkte der Linie richtige Ergebnisse liefert, für andere Punkte jedoch falsche Ergebnisse liefert, da die Minimierung der kleinsten Quadrate nur über einen eingeschränkten Wertebereich durchgeführt wurde. Die Verteilung der Werte

θ_v zeigt das folgende Histogramm. Es wurde für die erkannten Linien aus Abbildung 12 berechnet.

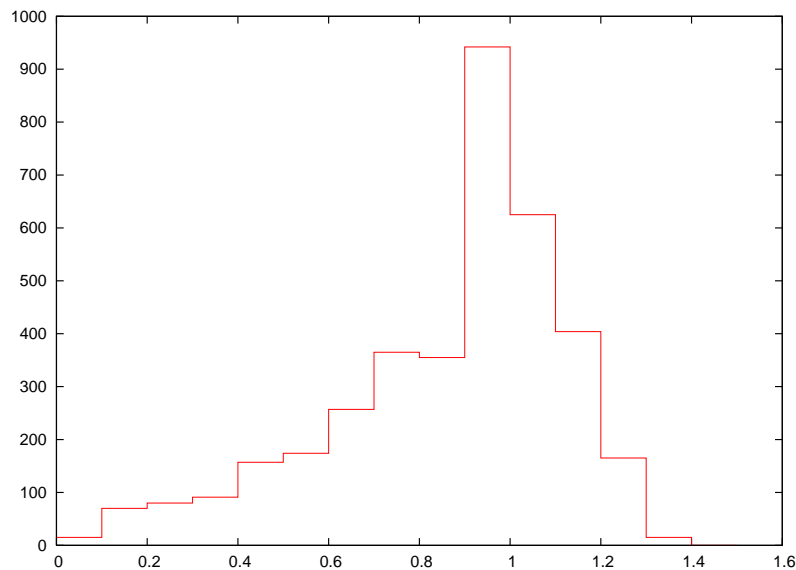


Abbildung 34: Histogramm zu Abbildung 12. Es zeigt die absolute Häufigkeitsverteilung der Abstände θ_v .

Desweiteren hat sich gezeigt, daß nicht vollständig erkannte Linien zu anders liegenden Ausgleichsgeraden führen, was das Ergebnis verschlechtert. Wird vor der Minimierung eine Skalierung der r_u derart durchgeführt, daß die Ausgleichsgeraden zu Tangenten der erkannten Linien werden, so entzerrt die errechnete Funktion das Bild wesentlich besser. Allerdings sind diese Ergebnisse noch nicht zufriedenstellend. Es ist daher notwendig, für Bilder von dem Lochkameramodell entsprechenden Kameras die Bestimmung der Entzerrungsfunktion iterativ durchzuführen. Bereits drei bis fünf Iterationen reichen aus, um gute Ergebnisse zu erhalten. Die verbleibende Krümmung der Linien im entzerrten Bild ist ein Maß für die Genauigkeit der berechneten Entzerrungsfunktion. Für Bilder von Kameras, die dem Kugelkameramodell entsprechen, sind die Ergebnisse auch ohne Iteration zufriedenstellend. Die Bilder von solchen Kameras zunächst in die Sicht einer Lochkamera umzurechnen, ist nicht lohnenswert. Die zum Bildrand hin abnehmende Auflösung im umgerechneten Bild erschwert das Erkennen der Linien. Paßt nicht das gesamte Bild, welches die Optik der Kamera liefert, auf den Bildsensor, so befinden sich schwarze Bereiche im umgerechneten Bild, so daß vom Programm Linien an den Kanten zu diesen schwarzen Flächen erkannt werden, was zu falschen Ergebnissen führt. Diese falsch erkannten Linien müssen entweder manuell gelöscht werden, bevor mit der Berechnung der Entzerrungsfunktion begonnen wird, oder der Blickwinkel muß eingeschränkt werden, wodurch das Ergebnis ebenfalls ungenau wird.

4.3 Ausblick

Das hier verwendete Kameramodell für Fischaugenkameras benutzt ein einziges optisches Zentrum. In diesem Punkt ist das Modell ungenau, denn Fischaugenkameras haben mehrere optische Zentren, welche sich auf der optischen Achse befinden. Es bietet sich an, das hier entwickelte Programm in diesem Sinne zu erweitern, um genauere Ergebnisse zu erzielen.

5 Literatur

- [1] Hynek Bakstein und Tomáš Pajdla. Omnivent Stereo-panoramas with Fish-eye Lens. Research Report (2001), Center for Machine Perception, Czech Technical University
- [2] Xianghua Ying und Zhanyi Hu. Distortion of fisheye lenses using a non-parametric imaging model. Asian Conference on Computer Vision (ACCV 2004), Jeju, Korea.
- [3] Janne Heikkilä und Olli Silvén. A four-step camera calibration procedure with implicit image correction. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'97), San Juan, Puerto Rico, p. 1106-1112.
- [4] Juho Kannala und Sami Brandt. A Generic camera calibration method for fish-eye lenses. 17th International Conference on Pattern Recognition (ICPR 2004), Cambridge, UK.
- [5] Wikipedia. <http://de.wikipedia.org/wiki/Kugelkoordinaten>